Geuvadis RNA-seq conference call minutes

September 16, 2011

Written by Tuuli Lappalainen (University of Geneva)

Sample set reminder from Geneva:

* We were informed at summer that the Iberian (IBS) cell lines will not be available until late 2011 at earliest. Waiting for them would have delayed the project by several months.
* We suggested then that these samples are replaced by 100 Yoruba. This was agreed on then, now mentioned again as a reminder
* **The final sample set is: 400 Europeans (CEU, GBR, FIN, TSI) and 100 Africans (YRI)**
* ECACC will send the cell banks of the 500 samples to Geuvadis partners in late October / early November

RNA extraction update from Geneva:

* the original plan was to have RNAs extracted at ECACC. Their two available protocols were tested in Geneva and Barcelona, and they turned out to be unsuitable for miRNA-seq.
* **All the RNA extractions will now be done in Geneva with Trizol**, that is known to work well. This will be done in October, with full randomization of the samples.
* RNAs of the randomly selected samples to be sequenced in each lab **will be sent out in late October**

Sample assignment:

* The number of samples sequenced at high coverage in each laboratory as agreed before
* **Geneva is willing to sequence all the 500 samples at lower coverage** (in addition to higher coverage of our share of the samples), to enable better control of batch effects. This was approved.

Discussion of library preps for mRNAseq:

* Two main options were discussed:

1. Do library preps only in Geneva, and send them to other labs for sequencing. However, in this case we would not be able to analyze differences between the labs caused by library preparation step
2. **Geneva will prep all the libraries, and the other labs will prep their own samples. Geneva will send each lab the libraries of the samples assigned to the labs, and each lab will sequence libraries both from their own lab and from Geneva**, half and half to obtain the desired coverage.

* Number two received most support. The final decision remains to be made, but for now the labs should prepare themselves to **start library preps in the end of October** when they receive the RNAs – preparing all the libraries in Geneva will take some time anyway.

**Small RNA seq** will be done according to the original plan, with **each lab preparing and sequencing their own samples**

Sequencing reagents and protocols

* We need to make an effort to use similar protocols as much as possible, at least the same version of the kit (the latest from Illumina). The decisions will need to be made soon so that people can order the reagents.
* **Geneva will suggest the kit to use for mRNA seq**
* **Barcelona will suggest the kit to use for small RNA**
* The level of indexing can be decided by each lab
* As a reminder: the previously agreed mRNA coverage is 15-20 M reads that map to coding sequence

Matthias from Kiel presented results from **strand-specific sequencing** with both Illumina and Solid protocols

* They are interested in working on this **within the Geuvadis framework**, which was agreed on.
* They have data from **5 samples now, but they are willing to do 10-15 more**
* They’re also interested in analyzing how genetic variation. Tuuli from Geneva offered to help with allele-specific expression analysis.
* Geneva has GRO-seq data from two 1000 genomes high-coverage trios, and it was agreed that these samples should be among those that will be analysed in the future. Geneva will see if it would be possible to run more GRO-seq for the 5 samples with strand-specific data

**Action items:**

* Geneva will inform everyone about the mRNA seq protocols to use
* Barcelona will inform everyone about the miRNA seq protocols to use
* Matthias will be in touch with people about how to proceed with the strand-specific project
* Final decision needs to be made about the distribution of libraries for mRNAseq

**Next call: late October** around the time when the RNAs are sent